“Econometric Analysis of realized Volatility: Evidence of Financial Crisis”

Ioannis Neokosmidis
PhD Candidate
Department of Economics
Aristotle University of Thessaloniki

Abstract. Financial series such as stock returns follow a different generating process from the relevant economic series. The key different between each other is that financial time series have some key features which cannot be captured by models such as ARMA. ARMA, which is referred as autoregressive moving-average, models consist a good approximation for economic series but not for financial series. In order to estimate financial time series we use the ARCH, autoregressive conditional heteroskedasticity, and GARCH, generalized autoregressive conditional heteroskedasticity, models. Moreover, we use six years data for four US stock indices such as, Dow Jones, NASDAQ, NYSE and S&P500, in order to analyse the volatility clustering and leverage effect. We conclude that the best fitted model for all our data is the EGARCH(1,1) in compare with an ARCH(6) or ARCH(4) and a GARCH(1,1). Additionally, we observed that the time periods between (28/07/2002-01/08/2003) and (11/08/2007-28/07/2008) are characterized by high volatility for all our series. In conclusion, we formulate and estimate multivariate volatility models, such as DVEC(1, 1), in order to show how the markets are linked by each other’s through time-varying covariance coefficients. The above methodology helps us to examine how the markets interact under the persistent of volatility effect. We use six years daily data from (26/3/2003) to (26/3/2009) in order to examine these interactions in S&P500, FTSE100 and DAX stock market indexes.

Keywords. Volatility; risk; ARCH; GARCH; EGARCH; Multivariate time series process.
1. INTRODUCTION

Financial world is based on the interaction between risk and returns. An investor must take some risks in order to achieve some level of wealth (rewards), but the general relationship between risk and rewards are not strictly analogy. Both risk and returns are in the future, so there is an expectation of loss a proportion of returns in balance with the risk that are taken. Financial econometricians express the exact relationship between risk and return as those risks which are calculated by the variance of the asset returns. It has been observed that volatility is not constant but it is changed over time, so it is larger for some period of time and smaller in other period of time. A standard approach in order to estimate the volatility is simply the sample standard deviation of returns in a time period, which called historical volatility. The main warning here is about the period of time. If we choose a short period of time for our sample of data we will get noisy results and if we select a long time horizon we will get results which will not be so relevant for resent measurements. So, historical volatility is not so reliable for further estimations and predictions. We need dynamic volatility models which will take the problem (warning) of time-varying volatility as a volatility that can be measured and not as a problem that must be corrected. Those models are the basic ARCH and GARCH models, which stand for autoregressive conditional heteroskedasticity and generalized autoregressive conditional heteroskedasticity and they were introduced by Engle (1982) and Bollerslev (1986) respectively.

Engle, R. (2007) argued that volatility is a fundamental factor in the global financial market. It is related with the risk that can be taken in order to have rewards. Risk and rewards are correlated each other but it is necessary to have a certain optimal behaviour in order to take risk, which can perform positive returns. So we choose a portfolio optimization position witch minimize the risk and maximize the rewards. Risk is determined by the variance of a portfolio in Markowitz (1952) theory for optimization. The same relationship between returns and variance can be sown in CAMP financial model which is introduced by Sharpe (1954). Moreover, risk can be determined very well by Black and Scholes (1972) model which is used in order to estimate the value of options in financial derivatives. The square route of variance is called volatility. Volatility is the standard deviation of the stock returns in a period of time. It is changing over time as it is presented by the analysts. We have different values of volatility in different time periods. Two basic types of volatility is the historical volatility and news volatility. The last is based on the element of information because every investor or risk manager would like to know if a small company will be developed in the future or not. Big companies give small volatility in contrast with small companies which give high volatility. So, if somebody knows that in a short period of time a company that is already small will be developed then he can arrange his investments in order to have arbitrage opportunities. Historical volatility, which is widely used, is estimated by historical data and it equals to the standard deviation of stock returns in a period of time. But if we get a short number of observations we will get noisy results and if we take a long series we will get smooth results which are not responding to the recent information. Historical volatility does not respond to that situation. ARCH models with their extensions come to fill this gap. ARCH (Autoregressive Conditional Heteroskedasticity) is introduced by Robert Engle in 1982 who won the Nobel price about that in 2003. ARCH volatility gives weights between the recent data and the data which are provided by information that happen a long time ago. The special feature of ARCH model is that it can calculate these weights based on historical data. There are lot extensions of ARCH models which describe non-linearity, asymmetry and long memory properties of volatility.
Dubofsky, D et al. (2003) argued that it is relevant to take the price of a call option as given and to formulate the variance of that option price. This variance or standard deviation is also referred as implied volatility (IV).

This article is organized as follows. In section 2, we introduce the formation of the fundamental volatility models which help us to examine the volatility clustering effect before and during the current financial crisis. In section 3, we give a description of our data and methodology in order to estimate the volatility processes which are analysed in the previous section. Finally, in section 4, we summarize the main findings.

2. FORMATION OF FUNDAMENTAL VOLATILITY PROCESSES

Let \( G \) be a subset of \( \mathcal{R} \) (\( G \subset \mathcal{R} \)). Then \( \forall t \in G \), the variable \( x_t(p) \) is a random variable (rv) which is defined in a probability space \( \mathcal{F} \) (\( \mathcal{F}: p \in \mathcal{F} \)). Then, a stochastic process \( \{x_t(p): \forall t \in G\} \) is referred as time series. If the elements of \( G \) are measured in discrete intervals then the time series process is a discrete time series, otherwise it is a continuous time series process with \( t \in [0, \infty) \). Generally a stochastic process can be characterized by its conditional distribution. If the distribution function of \( n \) variables are time independent then the data generating process is called strictly stationary. The previous assumption is not very practical as well as it is quite general in order to deal with it. So, we consider that it is remarkable to check the stationarity assumption only on some moments of the distribution function. If the mean, variance and covariance of a stochastic process are time independent then the time series model is characterized as covariance stationary or weakly stationary.

Another issue which is necessary to be clear, before we continue our analysis of financial time series, is the concept of linearity. Campbell, Lo, and MacKinlay (1997) define the linearity in a suitable way and also they underline an excellent structure of a time series model, which based on a Taylor series analysis, in order to show how a nonlinear stochastic process is formulated. Based on them, a general structure of a stochastic process is given by the following form:

\[
Y_t = f(u_t, u_{t-1}, u_{t-2}, ...)
\]

(I)

If we expand the (I) in a Taylor series around a point \( (u_t = 0) \) conditional on the information set \( \mathcal{F}_{t-1} : u_{t-i}, \forall i = 1, 2, ... \), we take the following parameterisation of (I):

\[
Y_t = g(u_{t-1}, u_{t-2}, ...) + u_t l(u_{t-1}, u_{t-2}, ...)
\]

(II)

\[
E_{t-1}(Y_t) = g(u_{t-1}, u_{t-2}, ...)
\]

(III)

\[
E_{t-1} \left[ (Y_t - E_{t-1}(Y_t))^2 \right] = l(u_{t-1}, u_{t-2}, ...)^2
\]

(IV)

We can see clearly from (III) and (IV) that models with nonlinear \( g(.) \) can be referred as nonlinear in mean and times series processes with nonlinear functions \( l(.)^2 \) are called nonlinear in variance. Basically, the financial time series models are all nonlinear in
variance or nonlinear both in variance and in mean. These models can capture the so called volatility clustering effect.

Generally, a financial analyst tries to approximate the relation between risk and return by using econometrics techniques for conditional variance which changes over time. This phenomenon is known as heteroskedasticity. This effect may give bias estimates for a sample of data and also affects the efficiency of statistical inference about the estimated coefficients of a relevant time series model. It is quite important to consider the dynamic variance as a factor that we have to deal with rather than a problem for our models. The first model which has this property is introduced by Engle (1982)\(^\text{1}\) and it is referred as ARCH (Autoregressive Conditional Heteroskedasticity). Moreover, there are a lot of extensions of ARCH that appeared in the relevant literature such as, ARCH-M, GARCH, EGARCH, TGARCH, GJR, AARCH, APARCH, FIGARCH, FIEGARCH, STARCH, SWARCH, GJR-GARCH, TARCH, MARCH, NARCH, SNPARCH, STUDENT-t-ARCH, but we focus only the mostly used symmetric models, ARCH, GARCH(1,1), ARCH-M and the most used asymmetric models, EGARCH and GJR which capture the asymmetric effects such as the leverage effect.

I. Univariate case

2.1 Autoregressive Conditional Heteroskedasticity (ARCH)

Let us suppose that our data are generated by an AR(p) process.

\[
Y_t = \rho_0 + \rho_1 Y_{t-1} + \cdots + \rho_p Y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2) \quad (2.1.1)
\]

With \(\sum_{i=1}^{p} \rho_i < 1\).

Since the unconditional mean of the above model is constant, there is not any useful meaning to use it in order to make forecasts. The optimal forecast for the (2.1.1) is given by the conditional mean of (2.1) which is given by:

\[
E(Y_t / Q_t) = \rho_0 + \rho_1 Y_{t-1} + \cdots + \rho_p Y_{t-p} \quad (2.2)
\]

Which \(Q_t\) represents the information set of our series, \(Y_t\).

The (2.2) represents the linear forecast of the mean of our series, but also we would like to find a process which forecasts the variance of the model.

We know that the residuals in (2.1) are white-noise, which means that the unconditional variance is constant and equals by \(\sigma^2\). The whole story is described by the unconditional variance of the errors, which may change over time. Let us suppose that the squares of errors are separately described by an AR(k) series such as below.

\[
\varepsilon_t^2 = b_0 + b_1 \varepsilon_{t-1}^2 + \cdots + b_k \varepsilon_{t-k}^2 + u_t \quad (2.3)
\]

---

Where $u_t$ is white-noise variables:

$$E(u_t) = 0 \text{ and } E(u_t u_{t-j}) = \sigma^2 \forall j = 0 \text{ or zero otherwise.}$$

As the equation (2.1), the linear forecast for (2.3) is given by the conditional mean such as:

$$E(\varepsilon_t^2 / Q_t) = b_0 + b_1 \varepsilon_{t-1}^2 + \cdots + b_k \varepsilon_{t-k}^2 \quad (2.4)$$

Where the information set, $Q_t$, includes that all the lagged values of $\varepsilon_t^2$: $Q_t = \{\varepsilon_{t-1}^2, \ldots, \varepsilon_{t-k}^2\}$.

Any representation such as (2.3) for every white-noise process is called an autoregressive conditional heteroskedastic process of order $k$ [ARCH(k)].

We see that the conditional variance of $\varepsilon_t^2$ is changed over time but the unconditional variance is constant and is given by:

$$\sigma^2 = E(\varepsilon_t^2) = \frac{b_0}{1 - b_1 - \cdots - b_k}$$

It is necessary to put some restrictions in our equations for conditional variance, in order to secure the possessiveness of volatility as any different result will be out of any meaning. So, this can be described by supposing that (2.4) is nonnegative and (2.3) is positive for every observation of $\varepsilon_t$. In order to take this result, we suppose that $b_i > 0 \forall i = 1, 2, \ldots, k$.

The equation (2.3) is stationary if $\sum_{i=1}^k b_i < 1$.

An ARCH model has many representations since the errors can be appeared in many different models such as an autoregression, an ARMA and the standard regression model. Basically, a linear representation such as (2.3) is not the most efficient since the model (2.1) and the models for the conditional variance are best estimated by MLE. So, a different approach is to represent the $\varepsilon_t$ as:

$$\varepsilon_t = \sqrt{h_t} \cdot z_t \quad (2.5)$$

With $z_t$ is an $(iid)$ process with zero mean and unity variance as while $h_t$ is described by:

$$h_t = b_0 + b_1 \varepsilon_{t-1}^2 + \cdots + b_k \varepsilon_{t-k}^2 \quad (2.6)$$

If any white-noise process is generated by the equations (2.5) and (2.6) then it is an ARCH of order $k$. Additionally, any linear forecast of (2.5) follows the same equations as (2.3) which is the conditional mean as represented by (2.4).

### 2.2 Estimation of ARCH

We can estimate an ARCH model by using MLE techniques. In order to explain how we can do this theoretically, we suppose that we have to estimate a regression model with ARCH disturbances such as: $Y_t = a_t X_{1t} + \varepsilon_t \quad (2.7)$
The disturbances follow the ARCH conditions of equations (2.5) and (2.6). Again we suppose that we have an information set, \( Q_t \), which includes all the lagged values of \((Y_t, X_{1t})\). With the above assumptions we conclude that the probability distribution for \( Y_t \) is given by:

\[
f(Y_t / X_{1t}, Q_t) = \frac{1}{\sqrt{2\pi h_t}} \exp \left[-\frac{(Y_t - a_1 X_{1t})^2}{2h_t}\right]
\]

(2.8)

Where, \( h_t = b_0 + b_1 \epsilon_{t-1}^2 \). The term \( \epsilon_{t-1}^2 \) equals the expression \((Y_{t-1} - a_1 X_{1t})^2\).

Our data vector is given by \( \delta = (a_1, b_o, b_1) \). We have to maximize the log likelihood function which given by,

\[
\mathcal{L}(\delta) = \sum_{t=1}^{T} \log f(Y_t / X_{1t}, Q_t, \delta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \log(h_t) - \frac{1}{2} \sum_{t=1}^{T} \frac{\epsilon_t^2}{h_t}
\]

(2.9)

2.3 Test for ARCH

Engle (1982) proposed a test for the time-varying variance (heteroskedasticity). This test involves some steps.

Step1: we estimate the regression model such as (2.7) by OLS by ignoring the heteroskedasticity.

Step2: obtain the residuals from the above regression and store it in our database.

Step3: we regress the squares residuals, \( \epsilon_t^2 \), on a constant and \( k \) of its own lagged values. If there are ARCH effects then the coefficients of lagged errors values must be insignificant from zero.

In model such as, \( \epsilon_t^2 = b_0 + b_1 \epsilon_{t-1}^2 + \cdots + b_m \epsilon_{t-k}^2 + u_t \), the estimated values of \((b_1, \ldots, b_k)\) must be other than zero.

If we conclude that there are ARCH effects then the coefficient of determination of the above model must be quite high. Under the null of no ARCH effect the test statistic, \( TR^2 \), follows a \( x^2 \) distribution. Here the T represents the magnitude of a sample of residuals and the \( R^2 \) the coefficient of determination. If the test statistic is sufficient low we conclude that there no ARCH effects and if the test statistic is high we conclude that there are ARCH effects.

2.4 Extensions of ARCH

GARCH:

The most widely used extension of ARCH is the GARCH model which is introduced by Bollerslev (1986). GARCH stands for \textit{generalized autoregressive conditional heteroskedasticity}.

Let us suppose that we have an ARCH specification for the residuals as:

\[ \varepsilon_t = \sqrt{h_t} \cdot z_t \]

With

\[ h_t = b_0 + b_1 \varepsilon_{t-1}^2 + \cdots + b_k \varepsilon_{t-k}^2 \quad \text{and} \quad z_t \sim iidN(0,1). \]

The above model for the conditional variance \( h_t \) can be rewritten as:

\[ h_t = b_0 + b_1 L^1 \varepsilon_t^2 + \cdots + b_k L^k \varepsilon_t^2 = b_0 + B(L)\varepsilon_t^2 \]

With \( B(L) = b_1 L^1 + \cdots + b_k L^k \), represents the lag operator.

If we generalized the above lag operator as a ratio of two lag operators such as,

\[ A(L) = \frac{B(L)}{C(L)}, \text{with} \ C(L) = 1 - c_1 L^1 - \cdots - c_m L^m \]

We conclude to a generalized representation of conditional variance as shown below.

\[ h_t = m + c_1 h_{t-1} + \cdots + c_m h_{t-m} + b_1 \varepsilon_{t-1}^2 + \cdots + b_k \varepsilon_{t-k}^2 \quad (2.10) \]

If a process \( \varepsilon_t \) is generated like (2.5) with a combination of (2.10) then we say that the \( \varepsilon_t \) is generated by a GARCH model, and we denote it like \( \varepsilon_t \sim GARCH(m,k) \).

We can see that if the ARCH is generated by an AR representation then the GARCH is generated by an ARMA representation and the polynomial operator \( A(L) \) is exactly the same as we move from an AR to an ARMA model. Furthermore, we can estimate a GARCH by MLE with the same philosophy as an ARCH and we can test for GARCH with the same method as an ARCH.

**IGARCH:**

Another extension of ARCH is the IGARCH model. A GARCH\((m,k)\) model is given by

\[ \varepsilon_t = \sqrt{h_t} \cdot z_t \quad \text{and} \]

\[ h_t = m + c_1 h_{t-1} + \cdots + c_m h_{t-m} + b_1 \varepsilon_{t-1}^2 + \cdots + b_k \varepsilon_{t-k}^2 . \]
The above process for conditional variance has unit root (stationary process) if the following condition is satisfied:
\[ \sum_{i=1}^{m} c_i + \sum_{j=1}^{k} b_j = 1 \] (2.11)

Any model which satisfies (2.5), (2.10) and (2.11) is referred as an integrated GARCH model and is denoted as IGARCH.

ARCH-M:

This model is introduced by Engle, Lilien and Robins (1987). We can say that it captures the relation between risk and return by considering that the mean of returns can be related with the variance of the returns. This relation is introduced by a regression model of the following form:
\[ r_t = \mu_t + \theta h_t + \varepsilon_t \] (2.12)

Where \( r_t \) is a series for returns, \( \mu_t \) represents the mean of returns and the term \( (\varepsilon_t) \) satisfies the conditions (2.5) and (2.6). The coefficient \( (\theta) \) capture the effect that higher risk, which is represented by the variance of \( \varepsilon_t \), gives higher returns \( (r_t) \).

EGARCH:

This model is introduced by Nelson (1991) and it is useful because capture some asymmetric effects which cannot be captured by the symmetric ARCH models like the above. The most interesting asymmetric effect is the leverage effect and it is related with the impact of news in volatility. More specifically, this effect is occurred when the volatility increases when actually the prices dropped (bad news) rather that when the prices are increased (good news) on similar level. The above effect cannot be explained by the ARCH and GARCH. So, Nelson (1991) proposed the following model for the conditional variance.

Let us suppose that the equation (2.5) is occurred. Then we have,
\[ \varepsilon_t = \sqrt{h_t} \cdot z_t \]
and,
\[ \log(h_t) = \delta + \sum_{i=1}^{\infty} \psi_i \{ |z_{t-i}| - E(|z_{t-i}|) + \theta z_{t-i} \} \] (2.13)
With \( z_t \sim iid N(0,1) \).
Any model with the characteristics of (2.13) is called exponential GARCH or EGARCH.

The asymmetric effect is expressed by the parameter \( \theta \) in the (2.13). If \( \theta = 0 \) then any positive sock has the same magnitude on volatility with any negative sock. If \(-1 < \theta < 0\) then a negative sock decreases volatility in a higher degree than any positive sock. When \( \theta < -1 \), any random negative sock increases the volatility while any positive random sock decreases the volatility.

GJR:

The GJR model also can capture the asymmetric effects of positive and negative random socks. GJR model for conditional variance is proposed by Glosten, Jagannathan and Runkle (1989) and can be described as following:

Again we suppose that

\[
\varepsilon_t = \sqrt{h_t} \cdot z_t, \text{ with } z_t \sim iid N(0,1).
\]

And

\[
\begin{aligned}
\varepsilon_t &= \delta + b_1 h_{t-1} + a_1 \varepsilon_{t-1}^2 + \theta \varepsilon_{t-1}^2 \cdot I_{t-1} \\

h_t &= \delta + b_1 h_{t-1} + a_1 \varepsilon_{t-1}^2 + \theta \varepsilon_{t-1}^2 \cdot I_{t-1}
\end{aligned}
\]

(2.14)

In the above expression, the factor \( I_{t-1} \) is a dummy variable. If \( \varepsilon_{t-1} \geq 0 \) then \( I_{t-1} = 1 \) and 0 otherwise. If we find a negative estimation of \( \theta \) then we expect to capture the leverage effect. Again we want the above expression for the variance to be positive. We can secure this if we put the restriction \( b_1 > 0 \) and \( a_1 + \theta > 0 \) for the parameters.

II. Multivariate case

Generally, it is observed that the markets are cointegrated each other which means that price movements of one market index can affect another market index. The fact of interrelated markets is a key factor in financial analysis and it can be captured statistically by multivariate time series models. Such models contain multiple return series of the cointegrated markets and the main propose is to analyse the effect of conditional covariance between them in order to examine the dynamic volatility processes among the multiple return series.

We consider that the returns of three stock indexes are modelled as the summation of a constant and the innovation of the series such as following:
\[ r_t = \mu + u_t \quad (2,15) \]

Where

\[ r_t = (r_{1,t}, r_{2,t}, r_{3,t})', \quad \mu = (\mu_1, \mu_2, \mu_3)', \quad u_t = (u_{1,t}, u_{2,t}, u_{3,t})' \]

The conditional covariance matrix of the innovation vector \( u_t \), given the information set \( \mathbb{Z}_{t-1} \), is defined as \( H_t = \text{Cov}(u_t | \mathbb{Z}_{t-1}) \). We apply a diagonal VEC model (DVEC) to our series for the volatility modelling, which refers to the time varying of \( H_t \). The DVEC(p,q) is defined as:

\[ H_t = C + \sum_{i=1}^{p} A_i \odot (u_{t-i}u_{t-i}') + \sum_{j=1}^{q} B_j \odot H_{t-j} \quad (2,16) \]

The (2, 16) is transformed to DVEC(1,1) as:

\[ H_t = C + A \odot (u_{t-1}u_{t-1}') + B \odot H_{t-1} \quad (2,17) \]

\( H_t = \begin{bmatrix} h_{11,t} & \cdot & \cdot \\ h_{21,t} & h_{22,t} & \cdot \\ h_{31,t} & h_{32,t} & h_{33,t} \end{bmatrix} \) is the covariance matrix, where \( h_{11,t}, h_{22,t}, h_{33,t} \) are the variance elements and the cross products are the covariance elements between each other. \( h_{21,t} \) expresses the time varying correlation between the elements (2,1) at time \( t \), \( h_{31,t} \) expresses the time varying correlation between the elements (3,1) at time \( t \) and \( h_{32,t} \) expresses the time varying correlation between (3,2) at time \( t \).

Matrix \( C \) contains the constant terms and it is given by:

\[ C = \begin{bmatrix} c_{11} & \cdot & \cdot \\ c_{21} & c_{22} & \cdot \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \], moreover the symmetric matrices (\( A, B \)) are formatted as \( C \) and contain the constant ARCH and GARCH coefficients respectively.

Analytically, the (2, 17) is written as:

\[ \begin{bmatrix} h_{11,t} & \cdot & \cdot \\ h_{21,t} & h_{22,t} & \cdot \\ h_{31,t} & h_{32,t} & h_{33,t} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdot & \cdot \\ c_{21} & c_{22} & \cdot \\ c_{31} & c_{32} & c_{33} \end{bmatrix} + \begin{bmatrix} a_{11} & \cdot & \cdot \\ a_{21} & a_{22} & \cdot \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \odot \begin{bmatrix} u_{1,t-1}^2 & \cdot & \cdot \\ u_{1,t-1}u_{2,t-1} & u_{2,t-1}^2 & \cdot \\ u_{3,t-1}u_{1,t-1} & u_{3,t-1}u_{2,t-1} & u_{3,t-1}^2 \end{bmatrix} + \begin{bmatrix} b_{11} & \cdot & \cdot \\ b_{21} & b_{22} & \cdot \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \odot \begin{bmatrix} h_{11,t-1} & \cdot & \cdot \\ h_{21,t-1} & h_{22,t-1} & \cdot \\ h_{31,t-1} & h_{32,t-1} & h_{33,t-1} \end{bmatrix} \quad (2,18) \]
If we solve the above equation, we conclude to six equations which three of them are the conditional variance equations and the other three are the conditional covariance equations.

So,

- **Variance equations**
  
  \[
  \begin{align*}
  h_{11,t} &= c_{11} + a_{11} u_{1,t-1}^2 + b_{11} h_{11,t-1} \\
  h_{22,t} &= c_{22} + a_{22} u_{2,t-1}^2 + b_{22} h_{22,t-1} \\
  h_{33,t} &= c_{33} + a_{33} u_{3,t-1}^2 + b_{33} h_{33,t-1}
  \end{align*}
  \]

- **Covariance equations**
  
  \[
  \begin{align*}
  h_{21,t} &= c_{21} + a_{21} u_{2,t-1} u_{1,t-1} + b_{21} h_{21,t-1} \\
  h_{31,t} &= c_{31} + a_{31} u_{3,t-1} u_{1,t-1} + b_{31} h_{31,t-1} \\
  h_{32,t} &= c_{32} + a_{32} u_{3,t-1} u_{2,t-1} + b_{32} h_{32,t-1}
  \end{align*}
  \]

Another multivariate model, which is appeared in literature by Engle and Kroner in 1995, is the BEKK model:

\[
H_t = CC' + \sum_{i=1}^p A_i (u_{t-i}' u_{t-i}) A_i' + \sum_{j=1}^q B_j H_{t-j} B_j'
\]

where \(C\) is a lower triangular matrix and \((A_i, B_j)\) are square matrices. The main warning of the above model is that there are \(K^2 (p + q) + \frac{K(K+1)}{2}\) parameters to be estimated which is not practical when \((p, q)\) are high.

Bollerslev (1990) proposed the constant-correlation model in order to deal with the above constrains (low estimated parameters and volatility equations). The main equation of this models is:

\[
H_t = C + \sum_{i=1}^p a_i a_{t-i}^2 + \sum_{j=1}^q b_j H_{t-j}
\]

Which in the case of GARCH(1,1) parameterisation is given by:

\[
H_t = C + a_1 a_{t-1}^2 + b_1 H_{t-1}
\]

The above multivariate volatility model is referred as constant conditional correlation (CCC) model and is the most efficient for large systems. It is easily estimated by two steps. The first
step contains estimation of all univariate models and then (2nd step) calculation of the

correlation between the standardize residuals. Moreover it contains easy likelihood functions

for the estimated elements.

Engle (2002a) proposed a parameterisation of the CCC model by allowing the correlations to
vary rather than to be constant over time (Dynamic conditional correlation-DCC- model).
The technique of this model is to formulate the squared elements (volatilities) in one set and
the cross products (correlations) in another set. So, the correlation elements can be modelled
by DCC as a separate process from volatilities.

3. DATA ANALYSIS AND RESULTS

III. Univariate case

Generally, we are interesting in volatility modelling of the US stock index market. We chose
to pick 6 years daily data prices in order to have a completed sequence for our series while
the relevant literature mentions that in order to have clear and reliable time series we have to
collect at least 4 years daily data. That is because daily series frequencies are more useful in
contrast with weekly or monthly frequencies due to louse crucial information. Finally, we
select to pick data from (28/7/2002) to (28/7/2008). These data are referred to closing prices
of US stock indices. While we are interesting to model the volatility of these indices, we have
to calculate the returns of that series in order to obtain the volatility clustering feature which
is appeared by plotting the time series sequences of returns. We did not select the examining
time period (28/7/2002-28/7/2008) randomly. We picked that time horizon for our analysis
because we did not want to include the period of 2001 which was stigmatized by the terrorist
attack in New York, while we know that every random sock affects our series in a bid level.
The key feature of volatility clustering is that shows the periods in which the market can be
characterized by low or high volatility. If the returns are shown to have large dispersion then
this period of time can be characterized by high volatility in the market and if the returns a
re appeared to have low dispersion, this period of time can be characterized by low volatility in
the market.

The prices and returns are determined as we have explained in the previous chapter. It is
necessary to use an econometric model for our returns in order to capture the volatility
clustering and the leverage effect. The basic econometric model for returns is the random
walk model (RW) which is given by,

\[ r_t = \delta + \epsilon_t \]  \hspace{1cm} (1)

The term (\( \delta \)) represents the mean value of returns. Another approach is to use an AR(1)
model for our returns sequence such as:

\[ r_t = \delta + \rho_1 r_{t-1} + \epsilon_t \]  \hspace{1cm} (2)
The above models cannot explain accuracy the financial features of returns. The key reason is that the financial time series much be analysed under the assumptions of stylized facts of financial returns which have already analysed in the previous chapter. In order to model the volatility effects we need ARCH and GARCH processes which also are generated by models such as (1)&(2).

We are going to describe our series by an ARCH(6), GARCH(1,1) and EGARCH models. So, an ARCH(6) model for returns is given by the following equation:

\[
\begin{align*}
  r_t &= \delta + \varepsilon_t, \\
  \varepsilon_t &= \sqrt{h_t} \cdot z_t \\
  h_t &= w + p_1 \varepsilon_{t-1}^2 + p_2 \varepsilon_{t-2}^2 + p_3 \varepsilon_{t-3}^2 + p_4 \varepsilon_{t-4}^2 + p_5 \varepsilon_{t-5}^2 + p_6 \varepsilon_{t-6}^2, z_t \sim \text{iid } N(0,1).
\end{align*}
\]

Similarly, a GARCH(1,1) representation for returns is estimated as:

\[
\begin{align*}
  r_t &= \delta + \varepsilon_t, \\
  \varepsilon_t &= \sqrt{h_t} \cdot z_t \\
  h_t &= w + p_1 \varepsilon_{t-1}^2 + b_1 h_{t-1}, z_t \sim \text{iid } N(0,1)
\end{align*}
\]

Finally, the EGARCH model for returns is captured as:

\[
\begin{align*}
  r_t &= \delta + \varepsilon_t, \\
  \varepsilon_t &= \sqrt{h_t} \cdot z_t \\
  \log(h_t) &= w + \sum_{i=1}^{\psi} \psi_i \{z_{t-i} - E(\{z_{t-i}\}) + \theta z_{t-i}\}, z_t \sim \text{iid } N(0,1)
\end{align*}
\]
We are going to base on the above models in order to analyse the results and moreover the volatility which is captured by them. We tested about ARCH effects with (3) but empirically we concluded that ARCH models with less lagged values of errors, such as ARCH(4), fitted better on some samples of data. In order to switch which model is better for a specific data set, e.g. Nyse, we based on the minimum AIC\(^3\) criterion.

The results of each volatility model for our four US indices are shown by tables (1) to (4). The numbers inside the brackets indicate the p-value for each estimated parameter. Generally the stationary assumption is hold for all models which is what we expect.

Analytically, we estimated an ARCH(6), GARCH(1,1) and EGARCH(1,1) models for the Down Jones index returns from (28/07/2002) to (28/7/2008) time period and the results are presented by table (1). All ARCH and GARCH coefficients are shown to be positive and significant at 1% level of significance, except the first lag for ARCH(6) model. Moreover, the sum of ARCH and GARCH elements \((p_1 + b_1)\) for the GARCH(1,1) are quite close to unity \((0.995984)\) which means that there are volatility shock effects in our series. The sum of ARCH and GARCH coefficients is very important because it shows the implication of a shock on returns, except that it is an indicator for the stationarity of our model. In order to understand the asymmetric effects, we can transform the EGARCH conditional variance for returns to the following form:

\[
\ln(h_t) = \delta + p_t \frac{\varepsilon_{t-1}}{h_{t-1}} + \theta \frac{\varepsilon_{t-1}}{h_{t-1}} + b_1 h_{t-1}
\]

The asymmetric effect such as the leverage effect is captured by the parameter \((\theta)\). In this case the value of this parameter \((-0.089691)\) is between \((-1)\) and \(0\) and it is significant at 1% level of significance, which means that a negative surprise shock affects in higher degree the volatility in contrast with a positive surprise shock. A negative surprise shock could be characterised by an unexpected drop in price and a positive surprise shock could be characterized by an unexpected rise in price.

Table 1: Volatility models for Dow Jones returns.

<table>
<thead>
<tr>
<th></th>
<th>ARCH(6)</th>
<th>DOWJONES GARCH(1,1)</th>
<th>EGARCH(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\delta)</td>
<td>0.000504659</td>
<td>0.00041973</td>
<td>0.000727822</td>
</tr>
<tr>
<td>(w)</td>
<td>2.60E-05</td>
<td>4.65E-07</td>
<td>0.00425087</td>
</tr>
<tr>
<td>(p_1)</td>
<td>0.0153726</td>
<td>0.0558836</td>
<td>0.102852</td>
</tr>
<tr>
<td>(p_2)</td>
<td>0.0909042</td>
<td>0.0009</td>
<td></td>
</tr>
<tr>
<td>(p_3)</td>
<td>0.101659</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>(p_4)</td>
<td>0.195116</td>
<td>0.0128851</td>
<td></td>
</tr>
<tr>
<td>(p_5)</td>
<td>0.215198</td>
<td>0.9401</td>
<td></td>
</tr>
<tr>
<td>(p_6)</td>
<td>0.128851</td>
<td>0.0999787</td>
<td></td>
</tr>
<tr>
<td>(b_1)</td>
<td>0.9401</td>
<td>0.995984</td>
<td></td>
</tr>
<tr>
<td>(\theta)</td>
<td>-0.089691</td>
<td>-0.089691</td>
<td></td>
</tr>
</tbody>
</table>

\(^3\) AIC stands for Akaike’s information criterion and is a measure of the goodness of fit of an estimated model. Generally, we choose that model, which has minimum AIC.
In order to choose the best model for Dow Jones data set, we based on the minimum AIC criterion. The most sufficient model is the EGARCH(1,1) with minimum value of (−6.75321211) in contrast with ARCH(6)-(−6.66556871) and GARCH(1,1)-(−6.7227259).

We follow the same analysis procedure for the NASDAQ index returns. The results are presented by table (2). We estimated an ARCH(6) model, because it has the minimum AIC in contrast with lower lagged ARCH models, a GARCH(1,1) and an EGARCH(1,1). We can see that all the ARCH and GARCH coefficients are positive and quite significant at 1% level of significance, except the first lagged value of the ARCH(6) model. Additionally, the sum of the ARCH and GARCH coefficients (0.998399) for the GARCH(1,1) model are below one and close to it which implies the stationarity of our model and moreover the persistence of volatility on our index’s returns. Since the value of parameter that captures the asymmetric effect for the EGARCH model (−0.0616451) is between (−1) and (0) and significant at 1% level of significance, we conclude that bad news (negative shocks) increases the volatility more that the good news (positive shocks).

Table 2: Volatility models for NASDAQ returns.

<table>
<thead>
<tr>
<th></th>
<th>ARCH(6)</th>
<th>NASDAQ GARCH(1,1)</th>
<th>EGARCH(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ</td>
<td>0.000721829 (0.007)</td>
<td>0.0004511127 (0.088)</td>
<td>0.000773026 (0.006)</td>
</tr>
<tr>
<td>w</td>
<td>4.42E-05 (0)</td>
<td>4.16E-07 (0.308)</td>
<td>-0.00453534 (0.787)</td>
</tr>
<tr>
<td>p1</td>
<td>5.79E-31 (1)</td>
<td>0.0389021 (0)</td>
<td>0.0681133 (0)</td>
</tr>
<tr>
<td>p2</td>
<td>0.10831 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p3</td>
<td>0.0972697 (0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p4</td>
<td>0.195399 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p5</td>
<td>0.184849 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p6</td>
<td>0.157866 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b1</td>
<td></td>
<td>0.959497 (0)</td>
<td>0.999182 (0)</td>
</tr>
<tr>
<td>θ</td>
<td></td>
<td></td>
<td>-0.0616451 (0)</td>
</tr>
<tr>
<td>(α+β)&lt;1</td>
<td>0.998399</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>-6.09742751</td>
<td>-5.13361795</td>
<td>-6.1525585</td>
</tr>
</tbody>
</table>

As we can see from the above table, the most sufficient model for our data set is the EGARCH model because it has the minimum AIC value (−6.1525585) in contrast with the others.

Volatility effects of the Dow Jones and NASDAQ indices can be described by the same model (EGARCH) and they represent the same asymmetric effect by the same way. The same results we can see for the S&P 500 and NYSE indices. Analytically, we estimate an ARCH(4) model, because it has the minimum AIC value comparing with the lower lagged ARCH(6), a GARCH(1,1) and an EGARCH(1,1). The results are presented by tables (3), (4). Again all the ARCH and GARCH coefficients are significant at 1% level of significance, except the first lag coefficient for the ARCH(4) model, and positive which imply the volatility shock on returns. The sum of ARCH and GARCH coefficients (0.996397), (0.994254) for S&P 500 and NYSE are close to one respectively which what we expect to be for the GARCH models. The coefficient (θ) is between (−1) and (0) for both EGACRH
models for both series and it implies the same asymmetric effect of that bad news increase the volatility more that good news.

<table>
<thead>
<tr>
<th>ARCH(1)</th>
<th>GARCH(1,1)</th>
<th>EGARCH(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta )</td>
<td>0.000631192 (0.001)</td>
<td>0.000611922 (0.001)</td>
</tr>
<tr>
<td>( \omega )</td>
<td>3.80E-05 (0)</td>
<td>6.43E-07 (0.002)</td>
</tr>
<tr>
<td>( \rho_1 )</td>
<td>0.5668468 (0.056)</td>
<td>0.9644086 (0.0)</td>
</tr>
<tr>
<td>( \rho_2 )</td>
<td>0.12345 (0)</td>
<td>0.7898883 (0)</td>
</tr>
<tr>
<td>( \rho_3 )</td>
<td>0.177397 (0)</td>
<td>0.92946 (0)</td>
</tr>
<tr>
<td>( \rho_4 )</td>
<td>0.92946 (0)</td>
<td>-0.0892391 (0)</td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.9942154 (0.0)</td>
<td>0.9942154 (0.0)</td>
</tr>
<tr>
<td>AIC</td>
<td>-6.4629851</td>
<td>-6.1308912</td>
</tr>
</tbody>
</table>

The next figures present the returns and the volatility clustering for each of the following series (DOW Jones, NASDAQ, NYSE and S&P500) respectively. We can observe that when we have excess returns for the financial time series, the volatility is high and if we have flatter returns the volatility is low. This description gives the volatility clustering effect. Generally, the periods of high volatility for all indices observed at the beginning of our estimation period and at the end of our estimation period, thus for time periods during (28/07/2002-01/08/2003) and (11/08/2007-28/07/2008).
IV. Multivariate case

We use six years daily data, (27/03/2003-27/03/2009), of three index markets in order to analyse the dynamic volatility process of the multiple returns series. The financial data are obtained by finance yahoo and they are referred to S&P500, FTSE100 and DAX indexes. We use the Eviews in order to estimate the DVEC(1,1) for multivariate volatility modelling. The results are as follows:

Table 1: Estimated coefficients for mean return equation and DVEC(1,1)
Based on the table 1, our equations about conditional variances and covariances are estimated as:

\[
\begin{align*}
    r_{DA\text{X},t} &= 0.000424 + u_{DA\text{X},t} \\
    r_{SP,t} &= 0.000951 + u_{SP,t} \\
    r_{FT\text{SE},t} &= 0.000402 + u_{FT\text{SE},t} \\
    h_{11,t} &= 1.07E - 06 + 0.069628u_{11,t-1}^2 + 0.921071h_{11,t-1} \\
    h_{22,t} &= 1.40E - 06 + 0.086118u_{22,t-1}^2 + 0.909612h_{22,t-1} \\
    h_{33,t} &= 6.23E - 07 + 0.118790u_{33,t-1}^2 + 0.888821h_{33,t-1} \\
    h_{21,t} &= 8.10E - 09 + 0.027091u_{21,t-1}u_{11,t-1} + 0.950562h_{21,t-1} \\
    h_{31,t} &= 4.12E - 08 + 0.032508u_{31,t-1}u_{11,t-1} + 0.909211h_{31,t-1} \\
    h_{32,t} &= 6.37E - 07 + 0.101354u_{32,t-1}u_{22,t-1} + 0.898372h_{32,t-1}
\end{align*}
\]

Figure 1 shows the time plot of returns for each series and figure 2 shows the estimated volatilities for continuously compounded returns for each index market. Moreover, figure 2 presents the time-varying correlations of DVEC(1,1) model for continuously compounded returns of the three index markets.

![Figure 1](image1)

**Figure 1:** Continuously compounded returns for series FTSE100, S&P500 and DAX. The time horizon is from (27/03/2003) to (27/03/2009).
As we expect, the time period after the second semester of 2007 is characterized by high volatility and correlation between the markets. The (S&P500) and (FTSE100) indexes are negative correlated but the (S&P500, DAX) are positive correlated at the time space which begins the financial crisis.

4. CONCLUSIONS

Returns have some special features that cannot be analysed by the basic time series models. That features are:

1. The distribution of the financial time series, such as stock index returns, have heavier tails in contrast with the normal distribution
2. The returns are uncorrelated for different time period
3. Large changes in returns are followed by large changes and moreover, small changes in returns are followed by small changes. So, there is a cluster for changes on returns.

Such characteristics as the above can be captured only by financial models for heteroskedasticity such as ARCH and GARCH and the family of them. The basic element of those models is that they consider the conditional time-varying variance as given and not as a problem that it is necessary to fix it. This effect is extremely important because real data are behaved like this. The effect which is accrued by the above is reported as volatility clustering.
We applied volatility models such as ARCH, GARCH and EGARCH on four US stock indices: Dow Jones, NYSE, NASDAQ and S&P 500. The time horizon for our data sample is six years from (28/07/2002) to (28/07/2008). The first two models capture the volatility clustering effect and the third model captures the leverage effect. We concluded that the EGARCH model is the best fitted process for all our sample of data, based on AIC minimum criterion. Additionally, it is observed that we have high volatility periods at the beginning and at the end of our estimation period for all stock indices which are the time spaces between (28/07/2002-01/08/2003) and (11/08/2007-28/07/2008). Finally, we formulate and estimate a multivariate volatility model, DVEC(1, 1), in order to show the volatility clustering and time varying covariances between three major stock markets (S&P500, FTSE100, DAX), which play an important role in the global financial world during the time interval of the current (crucial) financial crisis.

REFERENCES


